December 12, 2024

E-Filed

The Honorable Thomas S. Hixson
United States District Court for the Northern District of California
San Francisco Courthouse, Courtroom E – 15th Floor
450 Golden Gate Avenue
San Francisco, CA 94102

Re: *Kadrey, et al v. Meta Platforms, Inc.*; Case No. 23-cv-03417-VC (TSH)

Dear Judge Hixson:

  Plaintiffs in the above-captioned action and Defendant Meta Platforms, Inc. ("Meta") jointly submit this letter brief regarding the sufficiency of Meta's 30(b)(6) testimony with respect to Meta's efforts to acquire text data to train Llama, including Meta's alleged torrenting and seeding. The parties conferred at the time of Meta's 30(b)(6) testimony but were unable to reach a resolution.

I.   **PLAINTIFFS' STATEMENT**

Plaintiffs have recently learned that while engaging in widespread copyright infringement of Plaintiffs' and others' copyrighted works, Meta knowingly involved itself in the shadowy world of online hackers by "torrenting" these works.[1] In so doing, it not only *downloaded* hacked files with these copyrighted works, but was also "seeding," i.e. *sharing*, them with other data pirates globally. But the witnesses Meta *itself* chose to testify about these acts were woefully unprepared on this vital topic, and the Court should order Meta to provide testimony.

In his 30(b)(6) deposition on November 20, Meta Director of Product Management Michael Clark admitted that Meta acquired the LibGen shadow library via a data sharing protocol called "torrenting," using a platform known as LibTorrent. As Mr. Clark testified, LibTorrent required that any user seeking to download LibGen from its network—including here, Meta— must also *upload* that data *to allow other users to access and download it*. The upload requirement is termed "seeding."[2] Mr. Clark admitted that Meta had engaged in seeding LibGen data, i.e. copyrighted works, as a requirement for Meta to download from LibGen, Ex. A at 348:18-351:9. This admission means Meta knew it was facilitating copyright infringement by acting as a distribution point for other users of pirated books (via "seeding" on the LibTorrent platform).

Plaintiffs sought 30(b)(6) testimony concerning exactly which clients were used, what was seeded, to whom, and other technical details.[3] Yet no such details have been made available. All told, Plaintiffs sought 30(b)(6) testimony on Meta's efforts to acquire text data for use with Llama, including specifically torrenting and seeding, from three deponents that *Meta* designated on the topic—Mr. Clark on November 20, Nikolay Bashlykov on December 6, and Sy Choudhury on December 5. Meta needed to prepare at least one of these deponents on the topic, especially due to its on-the-record representation during Mr. Clark's testimony that Mr. Bashlykov would provide 30(b)(6) testimony on seeding. But none of the deponents were adequately prepared to testify.

A. No Meta Witnesses Were Adequately Prepared to Testify About Torrenting.

A "corporation has 'a duty to make a conscientious, good-faith effort to designate knowledgeable persons for Rule 30(b)(6) depositions and to prepare them to fully and unevasively answer questions about the designated subject matter.'" *Great Am. Ins. Co. of N.Y. v. Vegas Const.*, 251 F.R.D. 534, 539 (D. Nev. 2008). The Court may order additional time for such testimony when a party fails to adequately prepare its testifying witness. *See, e.g., Baird v. Blackrock Institutional Tr. Co.,* 2019 WL 365845, at *2 (N.D. Cal. Jan. 30, 2019).

---

[1] *See* Celia Kaechele, *Traditional Notions of Fair Play and Substantial Justice in the Age of Internet Interconnectivity: How Masking an IP Address Could Constitute Purposeful Availment*, 21 YALE J.L. & TECH. 59, 105 n.6 (2019) ("'Torrenting' is a popular form of peer-to-peer file sharing in which a single user downloads multiple pieces of a single file from several users simultaneously.").

[2] *See* Joseph M. Eno, *What Motivates Illegal File Sharing? Empirical and Theoretical Approaches*, 69 N.Y.U. ANN. SURV. AM. L. 587, 591 (2013) ("Colloquially, uploaders are known as 'seeders,' downloaders are 'downloaders,' and active client participants are 'peers.' Those who download more than they upload are called, pejoratively, 'leechers.'").

[3] Nor has such information come in the form of documents, responsive to at least RFP No. 85 (re: torrenting), concerning which bit torrent client Meta used, including how much it seeded, the list of "peers" it transmitted to and from, and application logs/tracking. Plaintiffs will brief this separately if needed.

Meta designated Mr. Clark to testify concerning Topic 1 ("Meta's access to, identification of, assembly, processing, and/or curation of datasets to train its Large Language Models").[4] Meta nevertheless informed Plaintiffs during Mr. Clark's deposition that a previously unidentified 30(b)(6) deponent, Mr. Bashlykov, would testify for Meta about the details of Meta's seeding. *See* Ex. A at 394:2-9 ("With respect to your questioning about seeding," Mr. Bashlykov is "prepared to talk about the seeding in more detail[.]"). Mr. Clark testified, among other things, that he did not "have those details" in response to Plaintiffs' question, "[H]ow would [Plaintiffs] find out how much data was shared" by Meta via seeding of LibGen. *Id.* at 357:19-22. Meta did not object to torrenting or seeding as irrelevant or outside the proper scope of 30(b)(6) testimony. Nor could Meta—Mr. Clark had already admitted Meta's knowledge that "access to" LibGen by torrenting (via LibTorrent) required seeding, i.e., sharing with others, the data. *Id.* at 348:18-351:9.

Then, when Plaintiffs tried to ask Mr. Bashlykov about it, he (and Meta's counsel) all but admitted that he was not prepared to answer Plaintiffs' questions either. When asked whether he agreed with Mr. Clark's testimony that "Meta has seeded but it tried to minimize the amount [of data] that it seeds," Mr. Bashlykov testified, "as I was mentioning before, **I don't have information to prove or disprove that seeding was happening**" at Meta. Ex. B at 62:12-19 (emphasis added). Further, when asked "Is Meta currently using torrent for any sourcing of its datasets?" he testified "**I don't have information to suggest that it is**." *Id.* at 60:6-11 (emphasis added). Mr. Bashlykov testified that his relevant preparation consisted only of reviewing a single document and otherwise only reviewing "the internal documentation which specifies which ports [are] open on the dev servers." *Id.* at 49:22-50:12. Meta's counsel also stated, "**I had neglected— I had forgotten about that point**" when Plaintiffs reminded Mr. Bashlykov that Meta represented he would testify about seeding "in more detail." *Id.* at 47:19-48:9 (emphasis added).

Mr. Choudhury also should have been prepared to testify about torrenting because Meta designated him for Topic 8. Topic 8 covered "Meta's efforts to acquire text data to train the Meta Language Models, including licensing efforts, if any," language Meta itself drafted and this Court adopted. Dkt. 252 at 3 ("Meta's proposal as to topic 8 is acceptable."); Dkt. 247 at 17. Yet Meta's counsel stated for the first time at Mr. Choudhury's deposition that Mr. Choudhury was not prepared to testify as to the entirety of Topic 8, only the subordinate phrase "including licensing efforts, if any." Ex. C at 8:9-11:7 (Meta counsel stating Mr. Choudhury was prepared only to testify on "commercial outreach including but not limited to licensing regarding acquiring text data").

Meta failed to prepare its witnesses on this topic. In so doing, Meta didn't just renege on its promise, but it violated Fed. R. Civ. P. 30(b)(6). The entire purpose of Rule 30(b)(6) depositions is to "avoid the 'bandying' by corporations where individual officers disclaim knowledge of facts clearly known to the corporation." *Great Am. Ins. Co. of N.Y.*, 251 F.R.D. at 538.

B. Request for Relief

Plaintiffs seek three hours of 30(b)(6) testimony of an adequately prepared witness on Meta's acquisition of text data, including torrenting and seeding, for use with Llama, during which Meta should be required to make available to the witness and Plaintiffs access to computers with identical source code/data and tools so they can use them in connection with the deposition.

---

[4] Torrenting and seeding also fall within Topic 2, another Topic on which Meta designated Mr. Clark and for which he similarly was not prepared. Topic 2 includes "Meta's access, copying, reproduction and/or use of the datasets[.]"

II.   META'S STATEMENT

The Court should not allow Plaintiffs to rewrite their 30(b)(6) topics at the close of discovery, particularly after exhausting their sixteen (16) hours of 30(b)(6) deposition time and taking over twenty (20) 30(b)(1) depositions. Meta's witnesses were fully prepared on each of the designated topics. To the extent the witnesses did not testify with the specificity Plaintiffs now demand, it is either because the right questions were not asked or because Plaintiffs seek detailed testimony on alleged "torrenting" and "seeding" that are beyond the scope of the 30(b)(6) topics.[5]

First, there is no claim in this case about alleged torrenting (a method of downloading files) or seeding (sharing files by uploading). The sole claim involves Meta's alleged copying of Plaintiffs' works to train the Llama models, and whether that is fair use. Although Plaintiffs have filed an eleventh-hour motion to amend the operative complaint to add claims about torrenting and seeding (Dkt. 300), those claims are not presently in the case, and, as detailed in Meta's opposition (Dkt. 329), Plaintiffs' new claims are untimely and inadequately pleaded. Plaintiffs' request for more 30(b)(6) testimony on these issues also contradicts Plaintiffs' representation that their amendment would ***not*** require any new discovery. *See* Dkt. 301-2 at 7 ("[T]hese claims are premised entirely on … deposition testimony already provided by Meta.").[6]

Second, Plaintiffs incorrectly contend that details about torrenting and seeding fall under either Topic 1 ("Meta's access to, identification of, assembly, processing, and/or curation of datasets to train its Large Language Models")[7] or Topic 2 ("Meta's access, copying, reproduction and/or use of the datasets").[8] A party propounding a 30(b)(6) deposition notice must make the noticed topics "reasonably particularized" and craft matters of examination with "painstaking specificity." *United States v. HVI Cat Canyon, Inc.,* 2016 WL 11683593, at *3, *7 (C.D. Cal. Oct. 26, 2016). Neither topic covers the testimony to which Plaintiffs now claim entitlement—*i.e.*, "which clients were used, what was seeded, to whom, and other technical details" of torrenting, or the answer to "[H]ow would [Plaintiffs] find out how much data was shared" by seeding. *Supra.*

"Seeding" involves uploading files to torrent sites. *See* Dkt. 301-2 at 1. Topics 1 and 2 do not reference uploading or sharing at all. Rather, they relate to Meta's access and use of datasets to train its Llama models. **Mr. Clark** testified for nearly ***9 hours*** as Meta's 30(b)(6) witness on these and other topics (in addition to 7 hours in his personal capacity about similar issues). His preparation involved reviewing numerous documents and discussions with nearly 20 Meta employees. *See* Ex. E at 9:20-10:20; Ex. F at 211:9-213:6. And to the extent the issue of seeding overlapped with the noticed topics as part of Meta's alleged access and use, Mr. Clark was prepared and testified about it. *See* Ex. F at 349:1-356:13 (testimony about discussions with Mr. Bashlykov, torrent, and seeding). More detailed testimony was neither requested nor required.

For **Mr. Bashlykov**, Plaintiffs do not rely on any specific 30(b)(6) topic as a basis for their

---

[5] Meta's Responses and Objections to Plaintiffs' Rule 30(b)(6) Notice are attached as Ex. D.

[6] Plaintiffs' request for three additional hours (nearly a 20% increase in the total number of hours) is another indicator that Plaintiffs are overreaching. As a comparison, spent three hours with Mr. Bashlykov on several much broader topics.

[7] Ex. D at 3. For Topic 1, Meta agreed to "designate one or more witnesses to testify generally regarding Meta's selection, downloading, assembly, and processing of text datasets used to pretrain the Meta Language Models." *Id.* at 4. Plaintiffs did not object to this narrowing.

[8] Ex. D at 4. For Topic 2, Meta agreed to a narrowing, which Plaintiff did not object to. *Id.* at 5.

demands. Instead, they cite a statement of Meta's counsel at Mr. Clark's deposition that questions about seeding could also be asked of Mr. Bashlykov. *See supra.* Notably, Meta's counsel did not promise a detailed discussion of seeding or agree that Plaintiffs could rewrite their 30(b)(6) topics. In fact, Meta's counsel *objected* that Plaintiffs' inquiry into specific details of seeding was beyond the scope of the 30(b)(6) topics. Ex. F at 356:14-357:17 ("I don't read the topics to include that nor I don't think reasonably put him on notice"). Meta's counsel then pointed out that Mr. Bashlykov—a natural person to question given that Mr. Clark had identified him as involved with the datasets at issue—was designated to testify about the datasets produced in this case under Topic 7, and thus would have further information. Ex. F at 394:2-9 ("With respect to your questioning about seeding, kind of without arguing about whether it's under 1 or 7, Mr. Bashlykov will be the 30(b)(6) witness for Topic 7 and he's prepared to talk about seeding in more detail.").[9]

Mr. Bashlykov was in fact prepared to testify about alleged seeding—to the extent such testimony was called for by Topics 1, 2, or 7. Plaintiffs simply chose not to question him on it and instead made a unilateral decision that he was not prepared. Ex. I at 150:14-22. Mr. Bashlykov was directly involved in the dataset download process that, according to Plaintiffs, resulted in the alleged seeding. *See, e.g.* Ex. H at 132:1-12 (personal involvement in torrenting); *Tradeshift, Inc. v. Buyerquest, Inc.,* 2021 WL 2222811, at *1 (N.D. Cal. June 2, 2021) ("The amount of preparation that is necessary for a Rule 30(b)(6) deposition depends on the nature of the topics and what information the witness already knows or can remember."). Mr. Bashlykov also spoke with several Meta employees and reviewed numerous documents, including internal documentation about the dev servers. Ex. I at 18:3-21:6; 49:22-50:18 ("I refreshed my knowledge of the work I did using the notebook document [tab 8] … . I looked into the internal documentation which specifies which ports open on the dev servers."). Rather than showing a lack of knowledge, as Plaintiffs allege above, Mr. Bashlykov confirmed that, based on this preparation, he was not aware of any evidence of seeding and testified about steps taken to limit or prevent seeding: "during torrenting of sci-mag, there were precautions in place to limit seeding." Ex. I at 53:14-55:21. He also confirmed that he was prepared to testify "about the measures that Meta took to prevent the seeding" of the dataset he downloaded via torrent. Ex. I at 152:3-13. Inexplicably given their present complaints, Plaintiffs' counsel chose ***not*** to ask Mr. Bashlykov detailed or follow-up questions on these subjects, including about the details of the dataset download process he spearheaded after counsel for Meta elicited that Mr. Bashlykov was prepared to discuss that subject. Ex. I at 153:3-4. Plaintiffs instead concluded he was not prepared and elected not to even ask him about several relevant documents, including documents about the download process. Ex. I at 150:14-22 ("I had several documents today that I had planned to talk with Mr. Bashlykov about that I'm not going to discuss with him because it is clear that he was not prepared to speak on this topic…."); *id.* at 152:3-8 (witness confirming that he was not asked questions about documents he had reviewed in preparation). That Plaintiffs did not obtain the testimony they hoped for from Mr. Bashlykov, and that Plaintiffs' counsel failed to fully question him on topics on which he was prepared to testify,

---

[9] Topic 7 covers "the origin and contents of the datasets Meta has produced in this case," Ex. D at 9. Meta agreed to designate a witness "to testify generally regarding the origin of the datasets Meta has produced in this case [identifying five datasets], how each dataset produced has been used by Meta in training the Meta Language Models, including, generally, whether Meta downloaded the datasets at multiple points in time." *Id.* at 10. Plaintiffs did not object to this narrowing ***or*** add Topics 1 or 2 to Mr. Bashlykov's 30(b)(6) deposition notice. Ex. G.

does not mean he was inadequately prepared.

Plaintiffs also fail to identify any 30(b)(6) topic for **Mr. Choudhury** that would require testimony on the details of seeding. He was designated on Topic 8: "Meta's efforts to acquire text data to train the Meta Language Models, including licensing efforts, if any." Ex. D at 11; *see* Dkt 252 at 3. This topic is not about Meta's acquisition of datasets more generally, but pertains to acquisition of text data *via licensing*, as confirmed by the parties' prior briefing on the proper scope of this topic. *See* Dkt. 247 at 17-22 (Plaintiffs, stating that Topic 8 "involves the commercial purposes of the uses (and the extant and developing licensing market for training data).") And even if Topic 8 were construed more broadly than licensing, it cannot reasonably be read to cover the alleged uploading or dissemination of information to third parties via seeding, as it pertains only to Meta's "efforts to *acquire*" data.

Finally, Plaintiffs' unexplained request for more 30(b)(6) time with "computers with identical source code/data and tools so they can use them in connection with the deposition" is vague and baseless. To the extent they are seeking access to the source code that Meta provided in this matter, Plaintiffs never sought access to the source code computer in either Mr. Clark or Mr. Bashlykov's deposition. In fact, Plaintiffs never referred to any source code in their questioning about seeding. Rather, Meta's source code was the subject of a *separate* deposition topic (Topic 10), about which Mr. Bashlykov testified in detail. Notably, Plaintiffs do not contend he was inadequately prepared on source code. To the extent Plaintiffs' request is seeking something else, no other "source code/data and tools" were requested during discovery, and such a request is improper here. Plaintiffs' motion should be denied.

III.   PLAINTIFFS' REPLY

Meta wrongly claims that torrenting plays no role in this case. *How* Meta acquired and copied Plaintiffs' works is clearly relevant to their direct infringement claim—just as Meta's brazen decision to torrent millions of pirated works despite recognizing the clear legal risks bears on Meta's willfulness. These issues are relevant regardless of the disposition of Plaintiffs' Motion for Leave to File Third Amended Consolidated Complaint. And Meta's characterization of torrenting as "alleged[ly]" occurring is baseless: it obviously did, and the exhibits to this letter show that.

Similarly, Meta's claim that torrenting and seeding fall outside all topics listed in the Rule 30(b)(6) notice is plainly untrue. First, "the reasonable particularity requirement of Rule 30(b)(6) cannot be used to limit what is asked of the designated witness at a deposition," and "the 30(b)(6) notice establishes the minimum about which the witness must be prepared to testify, not the maximum." *Louisiana Pacific Corp. v. Money Market 1 Institutional Inv. Dealer*, 285 F.R.D. 481, 486 (N.D. Cal. 2012). Second, several noticed topics expressly ask how Meta "access[ed]," "cop[ied]" and "assembl[ed]" its shadow library datasets, Ex. D at 3-4, and Meta did so, in large part, through torrenting and seeding—a method of downloading and copying pirated material that Mr. Bashlykov himself described in a case document as "legally not ok." Moreover, Mr. Clark's deposition testimony about "seeding" was made in reference to the sole document Meta's counsel included on the issue in Mr. Bashlykov's deposition binder—and Mr. Clark testified that he had even spoken with Mr. Bashlykov about this exact document. Meta's counsel expressly represented that Mr. Bashlykov could address these issues in more detail, and counsel therefore created a reasonable expectation that *someone* (and likely Mr. Bashlykov) would be prepared to testify on torrenting and seeding. Yet, somehow, Mr. Bashlykov had even less knowledge on the topic than Mr. Clark. This is corporate bandying: at least one Meta witness should have been prepared.

5

By: /s/ *Bobby Ghajar*

Bobby A. Ghajar
Colette A. Ghazarian
COOLEY LLP
1333 2nd Street, Suite 400
Santa Monica, CA 90401
Telephone: (310) 883-6400
Facsimile: (310) 883-6500
Email: bghajar@cooley.com
        cghazarian@cooley.com

Mark R. Weinstein
Elizabeth L. Stameshkin
COOLEY LLP
3175 Hanover Street
Palo Alto, CA 94304
Telephone: (650) 843-5000
Facsimile: (650) 849-7400
Email: mweinstein@cooley.com
        lstameshkin@cooley.com

Kathleen R. Hartnett
Judd D. Lauter
COOLEY LLP
3 Embarcadero Center, 20th Floor
San Francisco, CA 94111
Telephone: (415) 693-2071
Facsimile: (415) 693-2222
Email: khartnett@cooley.com
        jlauter@cooley.com

Phillip Morton
COOLEY LLP
1299 Pennsylvania Avenue, NW, Suite 700
Washington, DC 20004
Telephone: (202) 842-7800
Facsimile: (202) 842-7899
Email: pmorton@cooley.com

Angela L. Dunning
**CLEARY GOTTLIEB STEEN &
HAMILTON LLP**
1841 Page Mill Road, Suite 250
Palo Alto, CA 94304
Telephone: (650) 815-4121

By: /s/ *Maxwell V. Pritt*

BOIES SCHILLER FLEXNER LLP
David Boies (*pro hac vice*)
333 Main Street
Armonk, NY 10504
(914) 749-8200
dboies@bsfllp.com

Maxwell V. Pritt (SBN 253155)
Joshua M. Stein (SBN 298856)
44 Montgomery Street, 41st Floor
San Francisco, CA 94104
(415) 293-6800
mpritt@bsfllp.com
jstein@bsfllp.com

Jesse Panuccio (*pro hac vice*)
1401 New York Ave, NW
Washington, DC 20005
(202) 237-2727
jpanuccio@bsfllp.com

Joshua I. Schiller (SBN 330653)
David L. Simons (*pro hac vice*)
55 Hudson Yards, 20th Floor
New York, NY 10001
(914) 749-8200
dsimons@bsfllp.com
jischiller@bsfllp.com

Interim Lead Counsel for Plaintiffs

6

Facsimile: (650) 849-7400
Email: adunning@cgsh.com

Attorneys for Defendant Meta Platforms, Inc.

ATTESTATION PURSUANT TO CIVIL LOCAL RULE 5-1(h)

I hereby attest that I obtained concurrence in the filing of this document from each of the

other signatories. I declare under penalty of perjury that the foregoing is true and correct.

 Dated: December 12, 2024   BOIES SCHILLER FLEXNER LLP

                                                            /s/ Maxwell V. Pritt
                                                            Maxwell V. Pritt
                                                            Reed Forbush
                                                            Jay Schuffenhauer

                                                            Attorneys for Plaintiffs

8